

実践論文

日本人英語学習者コーパスにおける動詞—名詞コロケーション  
とコンビネーションVerb—Noun Collocations and Combinations in the Corpora of  
Japanese English Learners佐竹 由帆<sup>\*1</sup>  
Yoshiho Satake

Email: yoshiho.satake.sugitani@gmail.com

keywords: collocations, learner corpus, language transfer, verb, noun

英語学習者にとってコロケーションは重要だが難しく、コロケーションに関する学習者の問題は十分に研究されていない。本研究は、日本人英語学習者のコーパス英語動詞・名詞コロケーションの使用を分析する。学習者が動詞—名詞コロケーションを過剰使用または過少使用しているかどうか、よく使用される不自然な動詞—名詞の組み合わせがあるか、動詞—名詞コロケーションや不自然な組み合わせの使用は習熟度に応じて変化するかどうかについて、日本人英語学習者コーパス（JEFLC コーパスと NICE コーパス）と英語母語話者コーパス（BNC コーパスと COCA コーパス）を比較して調べた。調査の結果、すべてのレベルの学習者がネイティブスピーカーよりもはるかに少ないコロケーションを使用し、学習者が動詞と名詞の組み合わせを形成する場合に発生する母語転移には正しい用法につながるものもあり、不自然な動詞と名詞の組み合わせの割合は、習熟度とともに増加することが明らかになった。不自然な組み合わせに日本語から英語への単語の直訳の影響があるので、教師は学習者が知っている単語を量的に増やすだけでなく、様々な意味を持つ単語をどのように使うかについての質的な語彙指導を行うべきである。

Although it is accepted that collocations are not only necessary but also difficult for English learners, their problems regarding collocations have not been researched fully. The present study analyzes the use of English verb—noun collocations in the corpora of Japanese learners of English. Whether Japanese learners overuse or underuse English verb—noun collocations, the types of uncommon combinations of verb and noun they often use, and whether their ways of using collocations and combinations of verbs and nouns change according to proficiency levels are investigated by comparing their corpora—the Japanese EFL learner corpus (JEFLC) and the Nagoya Interlanguage Corpus of English (NICE)—with those of native English speakers—the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). The data indicate that learners at all levels use much fewer collocations than native speakers, and both positive and negative language transfers occur when learners form verb—noun combinations, and the proportion of uncommon verb—noun combinations increases with proficiency level. Since the literal translation of each word from Japanese to English may contribute to uncommon combinations, teachers should conduct not only quantitative vocabulary instruction to increase the number of words that learners know but also qualitative vocabulary instruction to teach learners how to use a word with various meanings.

---

\*1: 青山学院大学非常勤講師

## 1. Introduction

Understanding collocations constitutes a significant portion of understanding a word(1). According to Nation(2), “Language knowledge is collocational knowledge”. Native-like fluency can be developed by promoting collocational competence(2) (3). Learners sound unskilled while using the target language if they lack collocational knowledge(4). However, although collocations are accepted as not only necessary but also difficult for English learners, learners’ problems regarding collocations have not been thoroughly researched. More research is needed on how non-native speakers use collocations differently from native speakers.

## 2. Literature

Laufer and Waldman(6) investigates the use of English verb-noun collocations in the writing of Hebrews. They retrieved the 220 most frequently occurring nouns in the learner corpus and in the native-speaker corpus, extracted verb-noun collocations and compared verb-noun collocations used by Hebrews with verb-noun collocations used by native speakers. The data revealed that learners produced fewer collocations than native speakers and that interlingual errors continued to persist even at advanced levels.

Some studies contend that learners of English use fewer collocations than native speakers do(5) (6) (7). Informants’ first languages vary among these studies—different first languages(5), Hebrew or Arabic(6), Chinese or Swedish(7). English learners having a different first language may use collocations differently.

Concerning the relationship between collocations and parts of speech, Laufer and Waldman(6) and Nesselhauf(8) insist that it is difficult for learners to produce verb-noun collocations, while Siyanova and Schmitt(9) find that it is not difficult for learners to produce adjective-noun collocations. Non-native speakers’ use of verb-noun collocations would need to be investigated further.

There are conflicting arguments regarding

learners’ use of collocations. Zhang(10) states that there is a correlation between learners’ proficiency and the number and accuracy of collocations they use, while Howarth(5) denies this correlation. Laufer and Waldman(6) and Nesselhauf(8) identify some influence of learners’ first language on collocational errors, whereas Wang and Shaw(7) are skeptical of first language transfer. Further study is needed to clarify the relationship between learners and their use of collocations.

## 3. Research Questions

The present study analyzes the use of English verb-noun collocations, which are difficult for learners to produce correctly(6) (8), in the corpora of Japanese learners of English. The research questions are as follows:

1. Do Japanese learners of English overuse or underuse English verb-noun collocations?
2. What types of incorrect verb-noun combinations do Japanese learners of English frequently use?
3. Do Japanese English learners’ ways of using collocations and verb-noun combinations change with proficiency levels?

These questions are investigated by comparing the corpora of Japanese learners of English with those of native English speakers.

In this paper, the terms “verb-noun combination” and “verb-noun collocation” are defined as follows.

- verb-noun combination—every verb-noun unit
- verb-noun collocation—verb-noun unit that is found in at least one of two authoritative collocations dictionaries: *Oxford Collocations Dictionary for Students of English (OCD)* (11) or *Macmillan Collocations Dictionary (MCD)* (12). A set of verb-noun collocations is included in a set of verb-noun combinations. Although there are association measures for collocation extraction, such as the log-likelihood test and mutual information, I used collocations dictionaries to judge whether a verb-noun unit is a collocation, following a methodology by Laufer and Waldman(6), who conducted similar

studies with informants whose first languages were Hebrew or Arabic. I compared some of my results with theirs.

## 4. Method

### 4.1 Corpora

#### 4.1.1 Native speaker corpora

As the native speaker corpora, I primarily used the British National Corpus (BNC), and secondarily used the Corpus of Contemporary American English (COCA), because both are large-scale balanced corpora that can serve as prescriptive models for English learners. The BNC is an approximately 100-million-word corpus of written and spoken British English, while the COCA is an approximately 400-million-word corpus of written and spoken American English. Both have been compiled using various sources from the latter half of the 20<sup>th</sup> century. I used the COCA when referring to learners' use of verb-noun combinations that do not appear in the BNC.

#### 4.1.2 Learner corpora

As the learner corpora, I used the Japanese EFL Learner Corpus (JEFL) <sup>1</sup> and a non-native English speaker corpus in the Nagoya Interlanguage Corpus of English (NICE-NNS) <sup>2</sup> because the two corpora share common features. Both are corpora of topic-based timed essays in English written by Japanese learners of English without using a dictionary. The JEFL is an approximately 670,000-word corpus of essays in English by approximately 10,000 Japanese junior and senior high school students. Without using a dictionary, each student wrote an essay for 20 minutes, based on one of six topics. The NICE-NNS is an approximately 120,000-word corpus of essays in English by 342 Japanese university and graduate students. Without using a dictionary, each student wrote an essay for one hour, based on one of 11 topics.

<sup>1</sup> The junior and senior high school corpus data used in this paper are based on the Japanese EFL Learner (JEFL) Corpus, which was built with Yukio Tono as the director. The JEFL Corpus web search system (Shogakukan Corpus Network: SCN) was used as a search tool.

## 4.2 Procedure

Spelling errors in the learner corpora were corrected and counted when they were target verbs and nouns.

### 4.2.1 Ten high-frequency nouns in the BNC

First, a list of ten high-frequency nouns in the BNC was compiled by totaling the number of singular and plural forms for each noun. Although the tenth most frequent noun is "government," the eleventh highest, "life," was investigated because there is no verb-"government" combination in the JEFL. Therefore, the following ten high-frequency nouns in the BNC were investigated: "time," "year," "people," "way," "man," "day," "thing," "work," "child," and "life."

### 4.2.2 Comparison between the proportion of verb-noun collocations in verb-noun combinations in the BNC and learner corpora

I calculated the proportion of verb-noun collocations that the ten high-frequency nouns make in verb-noun combinations in the BNC and learner corpora. As I was concerned with a variety of verbs, I limited the nouns before investigating the verb-noun combinations. The first and second verbs on the left side of the nouns were investigated because an article or an adjective often precedes a noun, and because the probability of including non-verb-noun combinations needed to be minimized by limiting the number of target words. I conducted a lemma-based count of verbs. For example, both "went" and "going" were grouped into "go."

### 4.2.3 Verb-noun collocations that learners use with high frequency

Verb-noun collocations that learners use with high frequency were investigated. They

<sup>2</sup> The university corpus data used in this paper are based on the Nagoya Interlanguage Corpus of English: Non-Native Speaker data (NICE-NNS), which was built with Masatoshi Sugiuta as the director.

were qualitatively analyzed to identify why learners use them with high frequency.

#### 4.2.4 Verb–noun combinations used by learners that native speakers use with low frequency

Verb–noun combinations used by learners that were not found in the OCD(11), MCD(12), or BNC were counted and grouped into low-frequency combinations. Each combination was investigated in the original, and examples that are not verb–noun combinations were excluded. For example, in the case of the sentence, “I saw dream every day,” although “dream” and “day” are automatically classified as a verb–noun combination, I did not count it as a low frequency verb–noun combination because “day” is actually a part of “every day.” The combinations were qualitatively analyzed to identify why learners often use them rather than the verb–noun collocations that learners use with high frequency.

#### 4.2.5 Difference according to proficiency levels

A proportion of verb–noun collocations in verb–noun combinations in the learner corpora and low-frequency verb–noun combinations were analyzed in terms of their difference at three proficiency levels: junior high school, senior high school, and university.

### 5. Results

#### 5.1 Comparison between a Proportion of Verb–Noun Collocations in Verb–Noun Combinations in the BNC and Learner Corpora

**Table 1:** The number of verb–noun combinations and collocations and the proportions of verb–noun collocations in verb–noun combinations in BNC and in learner corpora

	combinations	collocations	collocations/ combinations
BNC	104033	43648	42.0%
Learner Corpora	3989	1115	28.0%

As Table 1 indicates, the percentage of verb–

noun collocations in verb–noun combinations in the BNC was obviously higher than that in the learner corpora. However, the degree of underuse of collocations varies depending on each noun. In the cases of “time,” “people,” and “day,” a proportion of verb–noun collocations in verb–noun combinations in the learner corpora was higher than in the BNC.

I then determine whether there is a difference according to proficiency levels. As Table 2 indicates, the percentage of verb–noun collocations in verb–noun combinations was highest in the senior high school corpus, and the chi-squared test data indicated that it was significantly higher than that in the junior high school corpus ( $\chi^2 = 5.89$ ,  $df = 1$ ,  $p = .0152$ ) and the university corpus ( $\chi^2 = 13.03$ ,  $df = 1$ ,  $p = .0003$ ). The proportions of verb–noun collocations in verb–noun combinations of the junior high school and university corpora were not significantly different ( $\chi^2 = 1.74$ ,  $df = 1$ ,  $p = .1869$ , ns).

**Table 2:** The number of verb–noun combinations and collocations and the proportions of verb–noun collocations in verb–noun combinations at three proficiency levels

	combinations	collocations	collocations/ combinations
Junior high	1240	331	26.7%
Senior high	1811	557	30.8%
University	938	227	24.2%

These results revealed that compared to native speakers, Japanese learners of English underuse verb–noun collocations at all proficiency levels. In addition, the proportion of verb–noun collocations in verb–noun combinations does not increase with learners’ proficiency level. Decreasing use of combinations by university students would be caused by their increasing use of more difficult verbs (see 5.4).

#### 5.2 Verb–Noun Collocations that Learners Use with High Frequency

**Table 3:** Verb-noun collocations that learners use with high frequency

	verbs	frequency	co-occurrence nouns (frequency)
1	have	579	time(497), day(42), child(15), life(10), work(9), way (6)
2	spend	82	time(58), life(12), day(9), year(3)
3	take	65	time(58), life(12), day(9), year(3)
4	do	58	thing(48), work(10)
5	enjoy	50	life(37), time(9), work(4)
6	meet	34	people(34)
7	make	31	thing(15), time(12), way(3), work(1)
8	lose	21	time(9), way(7), life(5)
9	live	17	life(17)

According to Table 3, learners used “have” with extremely high frequency. The verbs often co-occur with time ranked first, second, and third, and were used at all proficiency levels.

The following are some examples of verb-noun collocations that learners used with high frequency.

- I **have** no **time** to eat them in the morning. (Junior)
- We **had** two **days** for preparation. (Senior)
- Watching and playing sports is a good tool to **spend time** with them. (University)
- Likewise when you feel bored or sleepy at work, **take** some **time** to move. (University)
- Because after breakfast, we **do** many **things**. (Senior)
- So, I **enjoy** my **life**. (Junior)
- And he wanted to **meet people**. (Junior)

Apart from the fourth example, “take time,” the literal translation of phrases from Japanese in English appears to result in correct collocations because the meanings of these collocations are the same as the meanings of combinations of the literal translation of each word from Japanese to English. Learners could probably translate literally and form correct collocations if they remembered the meaning of each word that they learned.

### 5.3 Verb-Noun Combinations Used by Learners that Native Speakers Use with Low Frequency

**Table 4:** Verb-noun combinations by learners that native speakers use with low frequency

	verbs	frequency	co-occurrence nouns (frequency)	proficiency levels (frequency)
1	become	13	year(9), child(3), day(1)	junior(5), senior(7), university(1)
2	keep	10	life(10)	junior(3), senior(6), university(1)
3	sell	8	thing(8)	junior(4), senior(4)
3	study	8	thing(8)	senior(2), university(6)
5	experience	7	thing(6), work(1)	senior(3), university(4)
6	challenge	6	thing(6)	junior(1), university(5)
7	return	5	day(2), man(2), people(1)	junior(3), senior(2)
8	imagine	4	life(3), thing(1)	university(4)
8	select	4	thing(2), man(1), way(1)	senior(1), university(3)

Table 4 indicates frequent verb-noun combinations used by learners that native speakers use with low frequency.

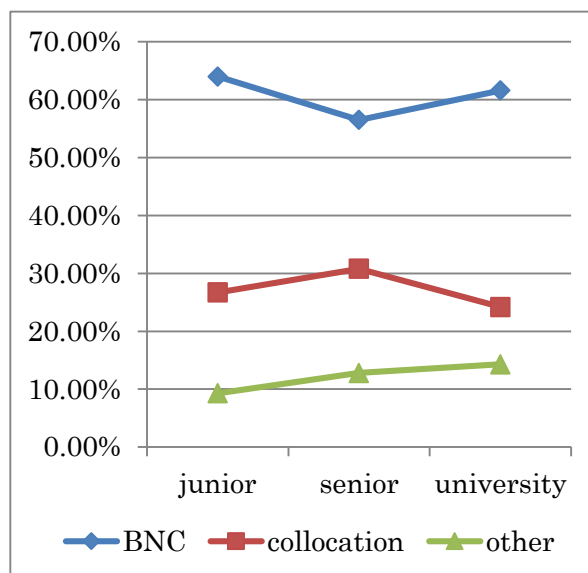
The nine verbs shared no word with the nine verbs that learners frequently use to form verb-noun collocations (see Tables 3 & 4). Learners at all levels used “become” and “keep” to form verb-noun combinations that native speakers use with low frequency. The following are examples of “keep life” and “become year,” which were most frequently used.

- He grew up, and **became** 80 **years** old. (Junior)

- He **became** three **years** old!! (Senior)
- Now she **became** 7 **years** old. (Senior)
- I can't **keep** my **life** without money. (Junior)
- Because, foods **keep** my **life**. (Senior)
- But people plays sports to feel relax, build up body, but not to **keep life**. (University)

“Become...years” is uncommon because it never appears in the BNC and appears only nine times in the COCA on this search condition. Although “Keep...life” appears 220 times in the COCA, it usually appears with an adjective (e.g., “keep your life organized,” “keep my life interesting”). Therefore, the above example of “keep...life” can be considered uncommon because it never appears in the BNC, and it appears only 13 times in the COCA in the same context as this search condition. Here it seems reasonable to suppose that “become...years” was used to mean “turn...years” and that “keep...life” was used to mean “maintain life” or “live.” In contrast to the case of verb–noun collocations by learners, the literal translation of phrases from Japanese to English appears to result in combinations that native speakers use with low frequency, because the combinations formed by the literal translation of each word from Japanese to English are not accurate. Thus, we see that whether learners form correct or unusual combinations is influenced by whether the literal translation of each word from Japanese to English forms a correct combination. Both positive and negative language transfers occur when learners form verb–noun combinations.

#### 5.4 How Learners Use Verb-Noun Combinations at Different Proficiency Levels



**Figure 1:** The proportion of verb–noun combinations at different proficiency levels

Figure 1 shows how learners use verb–noun combinations at different proficiency levels. The upper line indicates the percentage of verb–noun combinations that are not collocations but appear in the BNC, which comprise the largest percentage of all the combinations. The middle line indicates the percentage of verb–noun collocations, which comprise the second-largest percentage. The lower line indicates the percentage of verb–noun combinations that do not appear in the BNC, which comprise the smallest percentage. As mentioned before, we see that the proportion of verb–noun collocations in verb–noun combinations does not increase with learners’ proficiency level, because the percentage of collocations increases from junior high school to high school, and decreases from high school to university. We also see that the proportion of unusual combinations not appearing in the BNC increases as the proficiency level increases—this was similar to Laufer and Waldman(6), in which informants whose first language is Hebrew or Arabic use more unusual combinations as their proficiency level increases.

**Table 5:** High frequency verbs with which learners form unusual verb-noun combinations

	Junior high	Senior high	University
1	eat	eat	study
2	dream	become	<b>challenge</b>
3	become	enjoy	<b>experience</b>
4	practice	keep	imagine
5	sing	wonder	live
6	sleep	buy	learn
7	cry	dream	<b>memorize</b>
8	sell	like	<b>mention</b>
9	give	study	<b>select</b>
10	hope	think	start

Let us consider why unusual combinations increase with proficiency level. Table 5 indicates high-frequency verbs with which learners form verb-noun combinations that do not appear in the *OCD*(11), *MCD*(12), or BNC. The words in bold letters do not belong to the 1,000 highest-frequency words in authorized English textbooks in Japan(13), and are thus relatively difficult for learners. In contrast, the other words are among the 1,000 highest-frequency words(13), and are thus easier for learners than the words in bold letters. Relatively difficult words do not appear in junior and senior high school corpora, and are only present in the university corpus. From this perspective, the use of relatively difficult verbs may be considered to contribute to increasing unusual verb-noun combinations. Based on this assumption, increasing unusual verb-noun combinations is necessary at one stage of learners' development toward learning to use relatively difficult words.

## 6. Conclusions and Implications

### 6.1 Summary of Findings

Japanese learners of English underuse verb-noun collocations at all proficiency levels. This supports the results of previous studies (5)(6)(7). The proportion of verb-noun collocations in verb-noun combinations does not increase with learners' proficiency level. This supports the results of Howarth' study(5) and does not support the results of Zhang' study(10).

Whether learners form correct or unusual combinations is influenced by whether the literal translation of each word from Japanese to English forms a correct combination. This supports the results of the studies by Laufer and Waldman(6) and Nesselhauf(8) and does not support the results of Wang and Shaw(7).

The proportion of unusual verb-noun combinations that native speakers use with low frequency increases with proficiency level. This supports the results of Howarth' study(5) and does not support the results of Zhang' study(10). Increasing unusual verb-noun combinations is necessary at one stage of learners' development toward learning to use relatively difficult words.

### 6.2 Pedagogical Implications

Since verb-noun collocations were underused at all proficiency levels, it is necessary to conduct vocabulary instruction to promote learners' use of verb-noun collocations. Based on this study, vocabulary instruction to decrease learners' use of unusual combinations can be considered to contribute to relatively increasing their use of verb-noun collocations.

Since the literal translation of each word from Japanese to English may contribute to unusual combinations, teachers must make learners realize that literal translation is not always accurate. To raise learners' linguistic awareness, when a learner forms an unusual collocation, such as "became 80 years old," teachers should not only inform learners that a literal translation is not always accurate but also provide an example of a correct combination, such as "turned 80 years old." In addition, teachers should teach not only one primary meaning but also the various meanings of basic words. When a word reappears with a different meaning in a textbook, teachers should spend sufficient time on vocabulary instruction, because such a word is not always featured in a textbook. Furthermore, without sufficient vocabulary instruction, learners might not pay enough attention to a word's new meaning, assuming that they already know the word. If learners know, for example, that "turn" means not only "go round" but also "pass the age or time of,"

they can use that word in various ways and form correct combinations more easily. Because basic words often have various meanings, teachers should conduct not only quantitative vocabulary instruction to increase the number of words that learners know but also qualitative vocabulary instruction to teach learners how to use a word with various meanings.

### 6.3 Limitation and Suggestions for Further Research

One limitation of this study was that it was conducted on verb–noun combinations formed by only ten high-frequency nouns in the BNC. Thus, many verb–noun combinations formed by other basic and important nouns were not included in this study. To confirm the conclusions of this study, further research needs to be conducted on verb–noun combinations formed by other basic and important nouns.

learners of English and some implications for teaching”, *Applied linguistics*, Vol. 24(2), pp. 223–242(2003)

- (9) Siyanova, A. & Schmitt, N., “L2 learner production and processing of collocation: A multi-study perspective”, *The Canadian modern language review* Vol. 64(3), p. 429–458(2008)
- (10) Zhang, X., *English collocations and their effect on the writing of native and non-native college freshmen*, UMI(1993)
- (11) *Oxford Collocations Dictionary for Students of English*, OUP(2009)
- (12) *Macmillan Collocations Dictionary*, Macmillan(2010)
- (13) Shiomi, T., *The vocabulary of English language in the textbooks for junior and senior high schools*, Hokuseido(2002)

(2014年9月30日受付)  
(2014年12月3日採録)

### References

- (1) Zarei, A. A. & Mousavi, M., *Lexical collocations: Does feedback type matter?*, LAP LAMBERT Academic Publishing(2012)
- (2) Nation, I. S. P., *Learning vocabulary in another language*, Cambridge University Press, p. 318(2001)
- (3) Shin, D. & Nation, P., “Beyond single words: the most frequent collocations in spoken English”, *ELT journal* Vol. 62(4), pp. 339–348(2008)
- (4) Fahim, M. & Vaezi, R., “Investigating the effect of visually-enhanced input on the acquisition of lexical collocations by Iranian intermediate EFL learners: a case of verb-noun lexical collocations”, *Language teaching and research*, Vol. 2(3), pp. 552–569, (2011)
- (5) Howarth, P., *Phraseology in English academic writing: Some implications for language learning and dictionary making*, Niemeyer(1996)
- (6) Laufer, B. & Waldman, T., “Verb-noun collocation in second language writing: A corpus analysis of learners’ English”, *Language learning*, Vol. 61(2), pp. 647–672(2011)
- (7) Wang, Y. & Shaw, P., “Transfer and universality: Collocation use in advanced Chinese and Swedish learner English”, *ICAME journal*/Vol. 32, pp. 201–232 (2008)
- (8) Nesselhauf, N., “The use of collocations by advanced